Journal of Neuroscience Methods 322 (2019) 71–82

Contents lists available at ScienceDirect

Journal of Neuroscience Methods

journal homepage: www.elsevier.com/locate/jneumeth

EIQ: EEG based IQ test using wavelet packet transform and hierarchical extreme learning machine

Shashikant T. Kadam, Vineet M.N. Dhaimodker, Milind M. Patil, Damodar Reddy Edla*, Venkatanareshbabu kuppili

Department of Computer Science and Engineering, National Institute of Technology Goa, India

ARTICLE INFO

Keywords:
Electroencephalography
Hierarchical extreme learning machine
Brain computer interface
Wavelet packet transform
Human intelligence

ABSTRACT

Background: The use of electroencephalography has been perpetually incrementing and has numerous applications such as clinical and psychiatric studies, social interactions, brain computer interface etc. Intelligence has baffled us for centuries, and we have attempted to quantify using EEG signals.

New method: This paper aims at devising a novel non-invasive method of measuring human intelligence. A newly devised scoring scheme is used to ultimately generate a score for the subjects. Wavelet packet transform approach for feature extraction is applied to 5 channel EEG data. This approach uses db-8 as the mother wavelet. Hierarchical extreme learning machine is used for classification of the EEG signals.

Result: 80.00% training accuracy and 73.33% testing accuracy was measured for the classifier. The average sensitivity and specificity across all three classes was measured to be 0.8133 and 0.8923 respectively. An aggregate score was determined from the classification of EEG data. The power spectral analysis of the EEG data was conducted and regions of the brain responsible for various activities was confirmed. In the memory test, theta and beta bands exhibit high power, for arithmetic test, alpha and beta bands are strong, whereas in linguistic test, theta, alpha and beta bands are equally strong.

Comparison: The traditional IQ test determines intelligence indirectly, based on the score obtained from Wechsler test. In this paper an attempt is made to measure intelligence based on various brain activities – memory, arithmetic, linguistic.

Conclusion: A new method to measure intelligence using direct approach by classifying the EEG signals is proposed.

1. Introduction

Human have been trying to measure intelligence ever since this concept has been known. Many psychologists have tried to define intelligence. According to Alfred Binet intelligence is the judgment, otherwise called "good sense", "practical sense", "initiative", the ability of adapting one's self to circumstances (Binet, 1916). According to David Wechsler it is the aggregate or global capacity of the individual to act purposefully, to think rationally, and to deal effectively with his environment (Wechsler, 1944). But to measure the intelligence of a human being has always been a problem. Intelligence is one of the most highly debated and controversial subjects in psychology. Charles Spearman suggested that we have a comprehensive general intelligence that underlies all specific mental abilities, he called it the “g factor” (Reader et al., 2011). It is a psychometric construct that summarizes the correlations observed between an individual’s scores on a wide range of cognitive abilities and thus developed statistical procedure called factor analysis which was one of the first methods to correlate certain skills with others with the sole aim to measure intelligence (Reader et al., 2011). Further on, many tests to measure intelligence have been devised but the most popular and commonly used is the IQ test. However, none of these techniques directly utilize the brain signals. Brain is responsible for each and every action that a human being does. The brain transmits electrical impulses to the different parts of the body with the help of neural circuit. The neural circuit is a network of cells called neurons which are interconnected. The release of neurotransmitters such as dopamine, epinephrine, acetylcholine etc. trigger synaptic transmissions which lead to voltage change across the cell membrane (Lodish et al., 2000). The voltage change can be measured by placing electrodes on the scalp, this technique is known as

* Corresponding author.
E-mail addresses: shashikant.123kadam@gmail.com (S.T. Kadam), abcdvineet27@gmail.com (V.M.N. Dhaimodker), milindpatilmtg@gmail.com (M.M. Patil), dr.reddy@nitgoa.ac.in (D. Reddy Edla), venkatanaresh@nitgoa.ac.in (V. kuppili).

https://doi.org/10.1016/j.jneumeth.2019.04.008
Received 16 January 2019; Received in revised form 15 April 2019; Accepted 16 April 2019
Available online 22 April 2019
0165-0270/

Electroencephalography (EEG) (Niedermeyer and da Silva, 2004). Brain computer interface is then brought into action in order to transfer data to the computer. Data from the brain is collected in either invasive that is inserting electrodes on the cortical surface of the brain, or non-invasive that is placing the electrodes on the scalp (Abdulkader et al., 2015). BCI holds a number of limitations. The training process that is the preliminary stage is time consuming. Moreover, raw EEG data is predominantly non-linear, non-stationary and noise. A lot of techniques can be used to overcome these challenges. The main objective of the paper is to devise an efficient scoring scheme to measure the intelligence of the person using data collected directly from the brain. This diminishes any external manipulation by the subject during the test. The key intention is to uniquely classify the level of intelligence using brain signals in order to find the true mental ability of the subject. The intelligence of human beings has many aspects such as spatial ability, verbal comprehension, word fluency, perceptual speed, numeric ability, inductive reasoning and memory (Sternberg, 2019). The previous attempts of measurement of intelligence using EEG signals (Thatcher et al., 2016) (Ahmed et al., 2012) have made use of Wechsler test (Wechsler Test, 2019). In order to account for all the aspects of intelligence mentioned above, three tests namely memory, arithmetic and linguistic tests have been devised and an attempt to measure the human intelligence has been made. In the proposed test, we make use of the EEG signals that are generated during the mental activity stimulated by the tests to measure the intelligence of a person (Mustafa et al., 2013). Specifically, we target the following skills: memory, mathematical skills and linguistic skills. For noise reduction we have made use of band pass filter (Shenoi, 2006), for feature extraction wavelet packet transform and hierarchical extreme learning for multi class classification. The Section 2 defines the fundamental terminologies such as Wavelet Packet Transform and Hierarchical extreme learning machine. The proposed approach has been conferred in section 3. Result and analysis have been exhibited in section 4 followed by conclusion.

2. Terminology

2.1. Wavelet packet transform

The Wavelet Transform (Graps, 1995) (Grossmann and Morlet, 1984) of an original raw signal gives the time and frequency representation of the signal. After applying the transform, the data contained in the original signal is unaltered. For a given orthogonal function, a library of bases is generated. These bases are also called wavelet packet bases. These bases offer a unique method of coding signals, preserving total energy and reconstructing original features. Then, based on the requirements, the most suitable decomposition of the given raw signal is selected. Fig. 1 demonstrates the decomposition of the wave using wavelet packet transform. Wavelet packet transform (Singh and Ansari, 2016) (Ding, 2008) has been preferred over Fourier transform (Bochner and Chandrasekharan, 1949) in the proposed approach. This is because, one of the many applications of wavelet packet transform are the evaluation of “Best Basis”. Best Basis is a minimal representation of the data acquired on the basis of a particular cost function. "Best Basis” is considered suitable for feature extraction as it reduces noise and helps in data compression.

2.2. Hierarchical extreme learning machine

Hierarchical Extreme Learning Machine (H-ELM) (Tang et al., 2015) is an advanced model of Extreme Learning Machine (Huang et al., 2004) (Guang-Bin et al., 2006). It is an ELM extended to multiple layers. As opposed to general deep learning frameworks which use layer by layer backpropagation (Goodfellow et al., 2016), H-ELM has a multi-layered framework with two different feedforward learning phases:

1) Unsupervised phase, in which a sparse ELM autoencoder (Cao et al., 2016) is used. In this step, input weights are randomly initialised

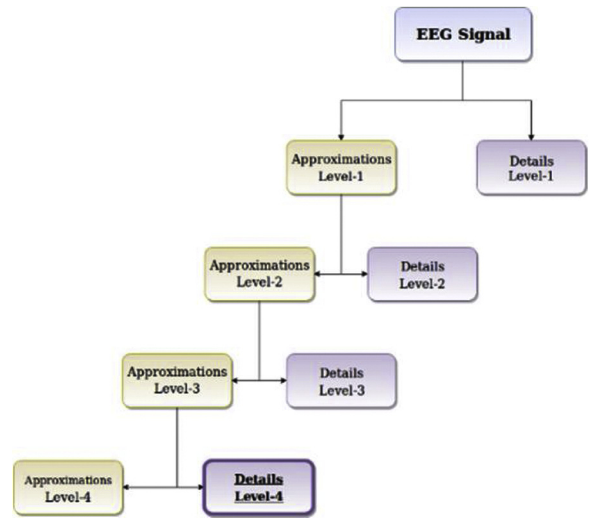

Fig. 1. Schematic Diagram of Wavelet Packet Transform.

and the sparse ELM autoencoder generates the hidden weights. The features are encoded across each layer in order to scatter them.

2) Supervised phase, in which the basic ELM algorithm is used to make final prediction. It makes use of Moore-Penrose inverse of a matrix (Moore, 1920) (also called pseudo-inverse of a matrix) in order to determine the weights in the final layer of the feedforward neural network. This method of supervised learning is much faster as compared to other machine learning methods as it involves only one matrix inverse and matrix multiplication. The architecture of H-ELM model is shown in Fig. 2.

H-ELM has been chosen for this model because of its fast learning ability and accurate predictions. Since it has multiple layers, it proves to be an excellent classifier for classification of signals. The H-ELM model used for the classification in this paper consists of 2 hidden layers since it suits well for relatively less data that we have. Using too many hidden layers may result in overfitting the training data (Leinweber, 2007). The result of decomposition of the EEG signals using wavelet transform at level 4 using DB-8 (Daubechies, 1992) as mother wavelet is used as the input layer. The value of stability factor $\left(\frac{1}{\lambda}\right)$ was set to 1.0×10^{-8}. The two hidden layers consist of 500 nodes each. Finally, the three outputs of the classifier were High, Average and Low.

3. Proposed approach

The determination of intelligence using EEG signal involves the following steps. First step is to setup the BCI device and the subject. Second step involves various tests and collection of EEG data. The acquired data is then preprocessed prior to feature extraction. Thereafter, wavelet packet transform is applied on the preprocessed data to obtain features for classification. These features are given as input to the classifier and the predictions are used to calculate the intelligence of the subject. A brief architecture of the entire process is displayed in Fig. 4 (EEG Head picture, 2019). A detailed flowchart is shown in Fig. 3.

First, it is made sure that subject does not have any mental medical history. He/she is familiarized with the procedure and written consent is taken from him/her. Next, the experiment is set up. The EEG cap is firmly placed on the subject’s scalp. The connections between the hardware are made and the electrodes are attached to the EEG cap at Fp1, Fp2, T3, T4 and Cz. Once the entire setup is ready, the subject starts to answer the tests. The test consists of 3 sections namely memory, arithmetic and linguistic. Each section includes 20 questions. While the subject is performing the test, his/her EEG signals are recorded. Subject is given 10 s to relax in between each question. Once the EEG signals are acquired, they are preprocessed by converting raw

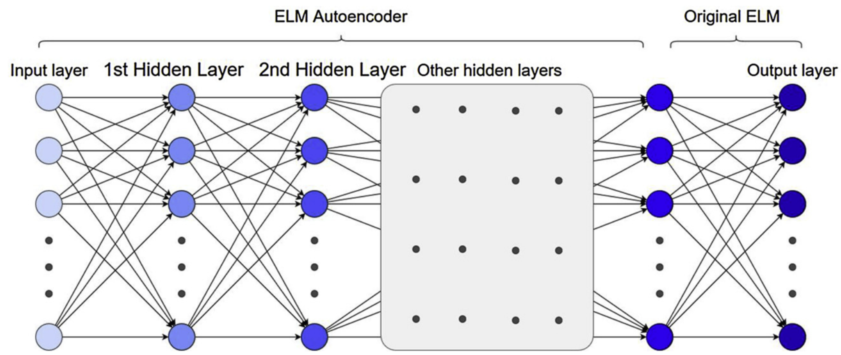

Fig. 2. Schematic Diagram of Hierarchical Extreme Learning Machine. The weights between input layer and first hidden layer and between successive layers are generated by the ELM autoencoder whereas the weights between last hidden layer and output layer are generated by the original ELM.

signals into amplified and digitized data with sample rate of 250 Hz. To remove noise, higher frequency components are removed using band pass filter in the range 0.3 Hz–30 Hz. Then, features are extracted from the data using Wavelet Packet Transform. Level 4 Wavelet Packet Transform is used and DB-8 is used as mother wavelet. These features are given as input to Hierarchical Extreme Learning Machine and data is classified into 3 categories using supervised learning. During the learning phase, the labels are decided using thresholds of the subject's performance in each test as explained in Table 1. The predicted classes are used to calculate the intelligence score of the subject.

A brief architecture of the entire process is displayed in Fig.4 which pictorially shows raw data collection from the subject, feature extraction to get cleaner and compressed data, process of classification of the subject based on this data and finally generating a score using the novel scoring scheme. A detailed flowchart shown in Fig. 3 describes in detail the entire process in 5 concrete modules which include a set of tasks. These modules are further explained in the sections that follow the figures.

3.1. Data acquisition

EEG data is collected using the Easy cap (Brain products, 2019) which consists of electrodes and electrode holders for connection. The position of the electrode holders on the cap is according to the 10/20 International system (Homan et al., 1987). Subjects are selected such that they have no history of brain injuries or any brain disorders. Every subject is asked to give written consent before taking the test. Subjects are instructed legitimately. The selected subjects are 18–22 years old. The two sections of the data acquisition process are as follows. 1) Setup: this section explains the preliminary task required to connect the hardware elements and run the software whereas 2) experiment: the experiment section explains the actual process recording the data.

3.2. Setup

The subject is seated comfortably in a silent room devoid of external noise or disturbance. The subject is asked to stay calm with minimal movement during the experiment to prevent superfluous modifications in the wave patterns due to physical movements. The EEG cap is placed on subjects' scalp. The electrodes are placed according to the 10–20 International system (American Electro-Encephalographic Society, 1994). All these electrodes are then attached to the EEG Easy cap. HiCl (High Chloride) gel is used before connecting the electrodes on the scalp to intensify the signals. The electrodes are also connected to the amplifier which is in turn connected to the computer. Fig. 7 is a representation of the experimental setup carried out. Brain Vision recorder and Brain Vision analyzer are the two software packages used (Brain products, 2019). Waves are recorded using the recorder whereas the pre-processing of signals is done using the analyzer.

Selection of electrodes is an important task involved in EEG data collection. It is advisable to select only the relevant electrodes for the desired application. Poor choice of electrodes may result in unwanted noise in the EEG signals. In this experiment, Fp1, Fp2, T3, T4, Cz electrode positions were chosen for data collection and Fz and P4 were used as ground and reference electrode respectively.

3.2.1. Fp1, Fp2 and cz

Fp1, Fp2 and Cz electrodes lie in the frontal lobe of the brain. Frontal lobe is responsible for decision making and consciousness. Besides, it is the region that contains maximum number of dopamine sensitive neurons. Cognitive processes related to attention, planning, reward, short-term memory and motivation are coordinated by the dopamine system (Imotions, 2016). Hence it was crucial to use the electrodes in the frontal lobe.

3.2.2. T3 and T4

T3 and T4 electrode positions are located in the temporal lobe of the brain. Temporal lobe is responsible for processing of sensory input using linguistics, visual memories etc. It is also associated with long term memory (Imotions, 2016). Hippocampus, responsible for memory encoding and retrieval is a central part of deep temporal structures. Besides, language processing, comprehension and production are conducted by left temporal cortex (Imotions, 2016). Fig. 5 (Image of brain, 2019) shows the relevant lobes of the brain and Fig. 6, (Image of 10-20 international system, 2019) shows the electrodes chosen from the 10–20 international system of electrodes.

3.3. Experiment

A test is prepared consisting of three sections: memory, arithmetic and linguistic. Each section consists of twenty questions. The test is administered via laptop or mobile device. For each question there are two stages: relax and perform. The subject is asked to relax for 10 s then perform for 20 s. This is repeated for each question in each section. The relax stage engages the subject to be calm whereas in the perform stage the subject has to answer the question of that particular section. Hence for a particular subject there are 60 recordings comprising of 1200 s. Appropriate relaxation time is given to the subject between any two sections.

3.4. Proposed test

The test includes three sections namely: Memory, Arithmetic and linguistic. Each section consists of 20 questions.

3.4.1. Memory test

The Memory test consists of two phases, the memorization phase and the recollection phase. In memorization phase the subject is shown

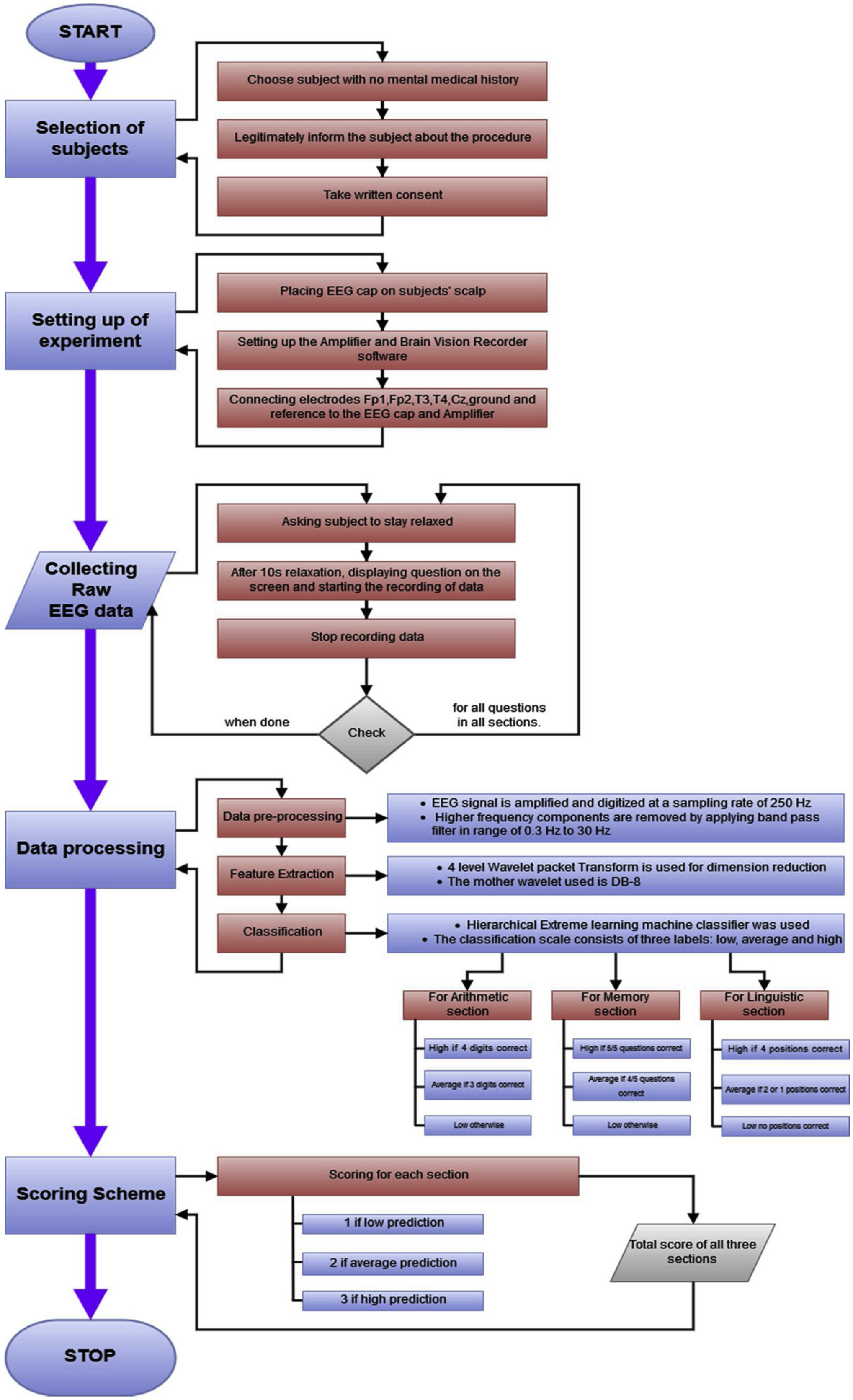

Fig. 3. Detailed flowchart of the proposed test.

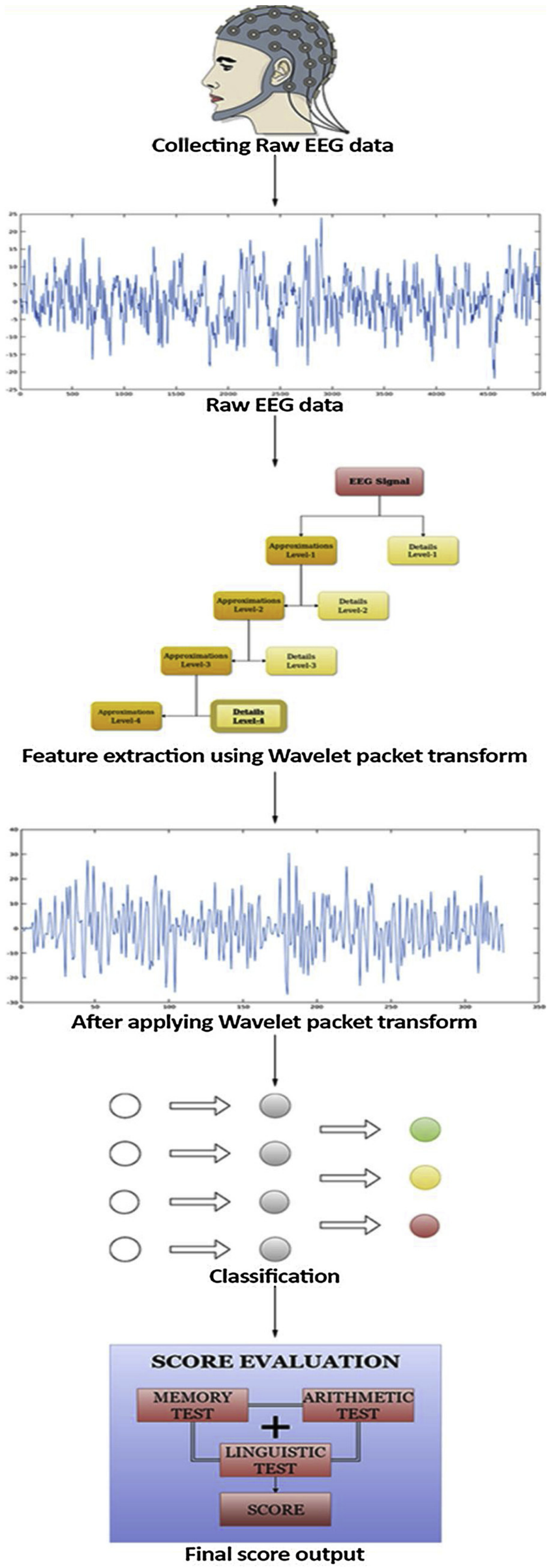

Fig. 4. Architecture of the proposed model.

a set of 5 three-digit numbers and he/she is asked to remember them within 10 s. Then the subject is allowed to relax for 20 s. After relaxation time, in the recollection phase, another 5 numbers are displayed sequentially to the subject and he/she is asked to respond yes or no based on whether each number was present in the previously displayed set of numbers. This was repeated for all 20 questions for all subjects. The EEG data is collected during both the memorization phase as well as recollection phase.

3.4.2. Arithmetic test

The questions in the arithmetic test are based on simple arithmetic operations, primarily consisting of addition and multiplication. A question is displayed on the screen and the subject is asked to solve it mentally under 20 s. The EEG data is acquired while the subject is mentally solving the question.

3.4.3. Linguistic test

In the linguistic test each question consists of a sentence broken down into 5–6 parts. The first and the last part of the sentence are given and remaining parts are jumbled. The subject is asked to rearrange the jumbled parts in order to form a grammatically correct sentence. Time limit for each question is 20 s. The EEG data is collected during these 20 s.

3.4.4. Signal pre-processing

Ahead of feature extraction and classification, preprocessing of raw EEG data is done. EEG signal is digitized and amplified at a sampling rate of 250 Hz. Analysis of the digitized data is done using Brain vision analyzer (Brain products, 2019). Band pass filter is applied in range of 0.3 Hz to 30 Hz to eliminate higher frequency components. Ultimately simulation is done using MATLAB (MATLAB, 2017) R2017a on Intel i5 processor with 8 GB RAM and Ubuntu 18.04 platform.

3.4.5. Feature extraction

In Orthogonal wavelet decomposition procedure, the general step is the process of depicting the raw signal as a set of overlapping wavelets. The overlapping is done on the basis of number of desired decomposition levels. Six level decomposition has been used as shown in Fig. 1. The details at the fourth level are then considered. For applying wavelet transform in the code, the wavedec (Wavedec (MATLAB), 2019) function in matlab is used. Wavedec function performs a multilevel wavelet analysis based on a specific wavelet. In this paper, Mother Wavelet used is DB-8. “DB" wavelets or (Daubechies (1992)) wavelets (Shen and Strang, 1998) are a set of orthogonal wavelets which define discrete wavelet transforms. In matlab, the wavedec function is called through the following syntax.

$$[c,\ l]\ =\ wavedec(x,\ n,\ wname)$$

where, wname-mother wavelet used, x-signal to be decomposed and n-number of decomposition levels.

For a given signal ‘x’, the function returns its wavelet decomposition at a level ‘n’ using the mother wavelet. In the returned wavelet decomposition, "c” is the wave decomposition vector which is returned as a real valued vector. "l” is the bookkeeping vector, which is returned as a vector of positive integers. The coefficients in the wavelet decomposition vector ‘c’ are parsed into levels using the bookkeeping vector ‘l’.

In this paper, detailed coefficients are obtained at level 4 using DB-8 as Mother Wavelet as mentioned earlier. These detailed coefficients are used as features for classification. The frequency range for these features is from 0.3 Hz to 30 Hz which include theta, alpha and beta bands of EEG signals (Imotions, 2016).

3.4.6. Classification

Classification is the most important step in the determination of intelligence using EEG data. An efficient classifier is the key for a

Table 1
Thresholds for labelling the datasets.

Test	Label	Criteria
Memory	High	5 out of 5 questions answered correctly
	Average	4 out of 5 questions answered correctly, 3 out of 5 questions answered correctly
	Low	2 out of 5 questions answered correctly, 1 out of 5 questions answered correctly, 0 out of 5 questions answered correctly
Arithmetic	High	4 digits correct
	Average	3 digits correct
	Low	2 digits correct, 1 digit correct, No digit correct
Linguistic	High	4 positions correct
	Average	2 positions correct, 1 position correct
	Low	No position correct

Fig. 5. Lobes targeted for the experiment (red region is frontal lobe and blue region is temporal lobe).

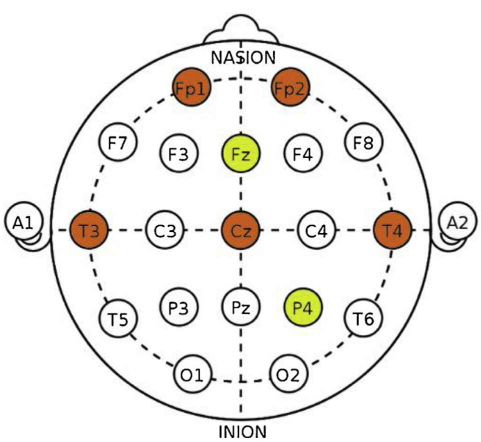

Fig. 6. The 10–20 International System of Electrodes (red electrodes are used in the experiment, Fz and P4 electrodes are used as ground and reference respectively throughout the experiment).

successful model. In this model, Hierarchical Extreme Learning machine (Tang et al., 2015) is implemented in a supervised learning manner with three classes namely, 1) High intelligence, 2) Average intelligence, 3) Low intelligence. The input layer is made up of concatenated features obtained from the wavelet packet transform on EEG data from the selected electrodes. Electrode Fp1, Fp2, T3 and T4 were used for memory test data, electrode Fp1, Fp2 and Cz were chosen for arithmetic test data and electrode T3 and T4 were chosen for linguistic test data. The number of hidden layer nodes was set to 500 for both hidden layers. The labels for training datasets were determined by the thresholds mentioned in Table 1.

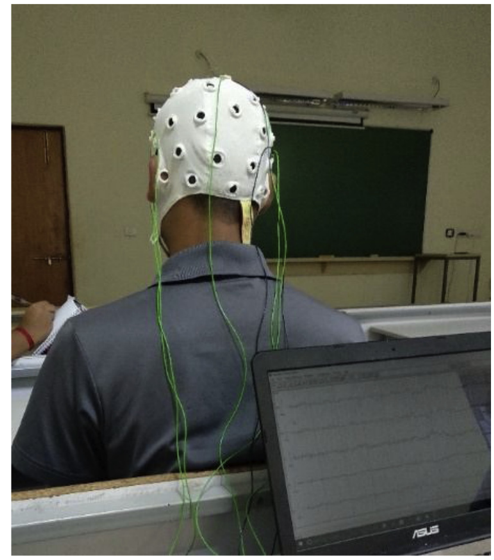

Fig. 7. A subject during experiment.

3.4.7. Hierarchical extreme learning machine

According to Jiexiong Tang, Chenwei Deng and Guang-Bin Huang in “Extreme Learning Machine for Multilayer Perceptron” (Tang et al., 2015), As shown in the section 2.2, The H-ELM is a feedforward neural network with two phases, unsupervised followed by supervised learning.

In the unsupervised phase, the ELM autoencoder tries to approximate the input data to make the reconstructed outputs appear similar to the inputs. The outputs are obtained using the following formula:

$$O_\beta = \underset{\beta}{argmin}\ \{\ ||H\beta - X||^2 + \ ||\beta||_{l1}\} \tag{1}$$

Where, X is the input data matrix, H represents the random mapping output, Beta represents the hidden layer weights which are obtained using a fast-iterative shrinkage-thresholding algorithm (FISTA) (Beck, 2009) and l_1 is penalty for the l_1 optimisation used in this step.

In the supervised phase, the original ELM algorithm is used with last hidden layer from the autoencoder as its input layer. Unlike Backpropagation, the weights of an ELM are generated in a single step. The ELM algorithm tries to minimize the following function (1):

$$||\beta||_u^{\sigma 1} + \lambda||H\beta - T||_v^{\sigma 2} \tag{2}$$

Where $\sigma 1 > 0, \sigma 2 > 0$, u, v = 0, (1/2), 1, 2, …, + ∞, H refers to the hidden layer output matrix (randomized matrix), T is the training data target matrix and Beta is the weights matrix. Thus, it not only minimises the training error but also minimises the norm of the weights.

The weights matrix Beta is obtained using Moore Penrose inverse given in (2).

$$\beta = H^\dagger T \tag{3}$$

Where T is target matrix, $H^\dagger$ is the Moore–Penrose generalized inverse of matrix H. The calculation of matrix $H^\dagger$ depends on whether the matrix HH^T is non-singular or H^TH is non-singular.

$$H^\dagger = \begin{cases} (H^TH)^{-1}H^T \text{ if } HH^T \text{ is singular} \\ H^T(HH^T)^{-1} \text{ if } H^TH \text{ is singular} \end{cases} \tag{4}$$

To further improve the stability of ELM, a positive value (1/λ) is added to the diagonal of H^TH or HH^T.

$$H^\dagger = \begin{cases} (\frac{1}{\lambda} + H^TH)^{-1}H^T \text{ if } HH^T \text{ is singular} \\ H^T(\frac{1}{\lambda} + HH^T)^{-1} \text{ if } H^TH \text{ is singular} \end{cases} \tag{5}$$

Table 2
Tabular form of accuracies.

Test	Dataset	Accuracy
Memory	Training	78.33%
	Testing	70.00%
Arithmetic	Training	85.00%
	Testing	80.00%
Linguistic	Training	76.66%
	Testing	70.00%

Therefore, Beta is given by (6) or (7).

$$\beta = \left(\frac{1}{\lambda} + H^T H\right)^{-1} H^T T \tag{6}$$

Or

$$\beta = H^T\left(\frac{1}{\lambda} + HH^T\right)^{-1} T \tag{7}$$

The corresponding outputs are given by (8) or (9).

$$\mathrm{f(x)} = \mathrm{h(x)}\beta = \mathrm{h(x)}(\frac{1}{\lambda} + H^T\mathrm{H})^{-1}H^T\mathrm{T} \tag{8}$$

Or

$$f(x) = h(x)\beta = h(x)H^T\left(\frac{1}{\lambda} + HH^T\right)^{-1} T \tag{9}$$

4. Results and analysis

In order to analyse the results of the proposed model, the acquired dataset is randomly divided into 4:1 ratio for training and testing datasets respectively. Training dataset is used to train the classifier and testing dataset is used solely for the purpose of testing the proposed model. After training the classifier, the training dataset is given as the input and the predictions are compared with the assigned labels in order to calculate training accuracy. Similarly, testing accuracy is obtained. The training and testing accuracies are mentioned in Table 2. SPSS software (IBM Corp, 2007) is used to calculate many of the various statistical results that follow.

$$Training\ accuracy = \frac{\text{correct predictions}}{\text{number of training samples}} \times 100$$

$$Testing\ accuracy = \frac{\text{correct predictions}}{\text{number of testing samples}} \times 100$$

Fig. 8 is a bar graph comparing training accuracies and testing accuracies. Table 3 gives the confusion matrix (Powers, 2011) of all three tests for all the subjects.

$$Sensitivity\ (Low) = \frac{True\ Positive}{True\ Positive + False\ Negative} = \frac{55}{55 + 13} = 0.\ 8088$$

$$Specificity\ (Low) = \frac{True\ Negative}{True\ Negative + False\ Positive} = \frac{137}{137 + 19} = 0.8782$$

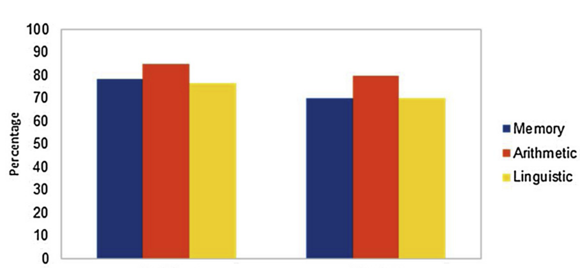

Fig. 8. Bar graph depicting the training and testing accuracies of the classifier.

Table 3
Confusion Matrix.

Predicted Class					
True Class		Low	Average	High	Total
	Low	55	6	7	68
	Average	14	80	14	108
	High	5	2	57	64
	Total	74	88	78	

$$Sensitivity\ (Average) = \frac{True\ Positive}{True\ Positive + False\ Negative} = \frac{80}{80 + 28} = 0.7407$$

$$Specificity\ (Average) = \frac{True\ Negative}{True\ Negative + False\ Positive} = \frac{112}{112 + 8} = 0.9333$$

$$Sensitivity\ (High) = \frac{True\ Positive}{True\ Positive + False\ Negative} = \frac{57}{57 + 7} = 0.8906$$

$$Specificity\ (High) = \frac{True\ Negative}{True\ Negative + False\ Positive} = \frac{135}{135 + 21} = 0.8654$$

$$Average\ Sensitivity = \frac{Sensitivity\ (Low) + Sensitivity\ (Average) + Sensitivity\ (High)}{3} = \frac{0.8088 + 0.7407 + 0.8906}{3} = 0.8133$$

$$Average\ Specificity = \frac{Specificity\ (Low) + Specificity\ (Average) + Specificity\ (High)}{3} = \frac{0.8782 + 0.9333 + 0.8654}{3} = 0.8923$$

4.1. Scoring scheme

For the proposed model, a new scoring scheme has been designed for the better measurement of intelligence of the subjects. The scoring has been solely determined by the predictions made by the classifier. The score for each question is obtained according to Table 4. The total score for each subject is calculated as follows:

$$Memory\ Intelligence\ Score = \sum_{i=1}^{20} Score\ for\ i^{th}\ question\ in\ memory\ test$$

$$Arithmetic\ Intelligence\ Score = \sum_{i=1}^{20} Score\ for\ i^{th}\ question\ in\ arithmetic\ test$$

$$Linguistic\ Intelligence\ Score = \sum_{i=1}^{20} Score\ for\ i^{th}\ question\ in\ linguistic\ test$$

$$Total\ Intelligence\ Score = Memory\ Intelligence\ Score + Arithmetic\ Intelligence\ Score + Linguistic\ Intelligence\ Score$$

The Fig. 9 maps each subject to one of the 5 categories based on the

Table 4
Score for each prediction.

Prediction	Score
High	3
Average	2
Low	1

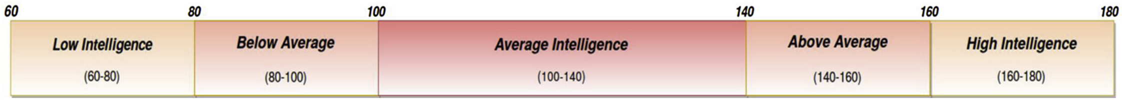

Fig. 9. Block diagram of different categories of intelligence measure based on total score.

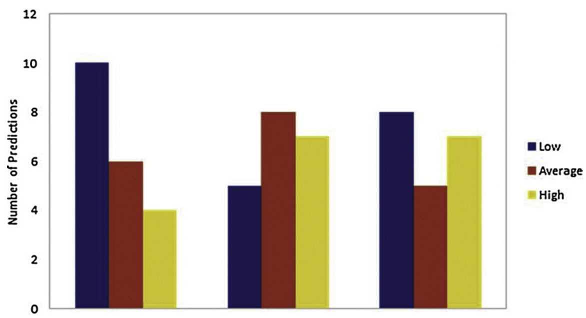

Fig. 10. Bar graphs representing the number of high, average and low predictions by the classifier in Memory test, Arithmetic test and Linguistic test for Subject 1.

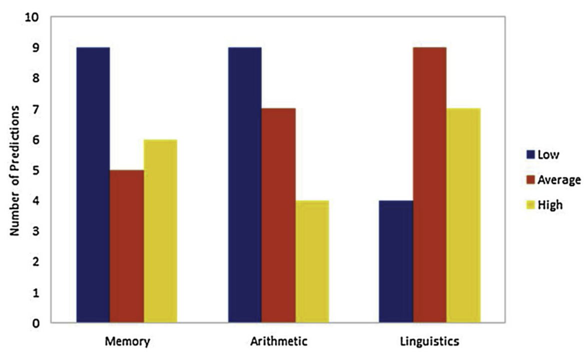

Fig. 11. Bar graphs representing the number of high, average and low predictions by the classifier in Memory test, Arithmetic test and Linguistic test for Subject 2.

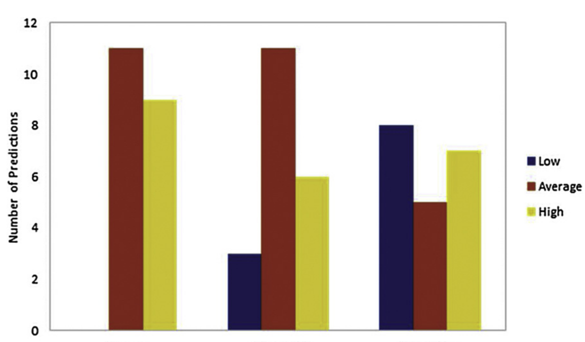

Fig. 12. Bar graphs representing the number of high, average and low predictions by the classifier in Memory test, Arithmetic test and Linguistic test for Subject 3.

total score of all the three tests.

4.2. Subject wise analysis

The following tables and bar graphs (Figs. 10–13) corresponding to the tables represent the subject wise analysis of data acquired. Each table consists of 20 rows signifying the number of questions and the columns represent the labels, predictions and scores for each question. The training accuracy for memory test is found to be 78.33%. Whereas the testing accuracy for the same is slightly lower, 70%. The training

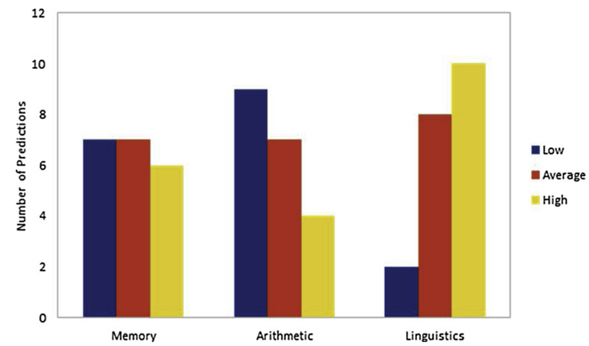

Fig. 13. Bar graphs representing the number of high, average and low predictions by the classifier in Memory test, Arithmetic test and Linguistic test for Subject 4.

accuracy for arithmetic test is 85% and testing accuracy for it is measured to be 80%. For linguistic test training accuracy is 76.66 and testing accuracy is same as that of memory test, 70%. For each prediction of the classifier, different scores are assigned. 3 score is assigned for high prediction. For average it is 2 score and for low, 1 score is assigned. The aggregate score is considered in determining the intelligence.

Table 5 gives the detailed result of subject 1. In the memory test, subject 1 answered 4 questions perfectly, 9 questions with 4 out of 5 numbers correct and remaining 7 were labelled as low. However, the predictions made by the classifier using the EEG data gives 4 questions as high, 6 questions as average and 10 questions as low. The total score for memory test is calculated to be 34. In arithmetic test, the subject's performance was as follows. 3 questions were labelled as high, 10 were labelled as average and 7 were labelled as low. The classifier gives 7 high, 8 average and 5 low predictions. Subject's arithmetic score is found to be 42. In linguistic test, 6 questions were labelled as high, 8 were labelled as average and 6 were labelled as low. Classification results in 7 questions predicted as high, 5 as average and 8 as low. The linguistic score is 39, giving a total of 115 intelligence score. This score maps subject 1 to average intelligence using Fig. 9.

The results of subject 2 are given in Table 6. Subject 2 had 5 questions labelled as high, 9 questions labelled as average and 6 questions labelled as low. The classifier predictions are as follows. 6 questions are predicted as high, 5 questions are predicted as average and 9 are predicted as low. The total score for memory test is 37. In arithmetic test, subject 2 answered 4 questions perfectly, 10 questions with 3 digits correct and remaining 6 questions were labelled as low. The predictions made by classifier are as follows. 4 questions are predicted as high, 7 are predicted as average and 9 are predicted as low. Subject's arithmetic score comes to a total of 35. Subject had 6 questions labelled as high, 11 labelled as average and 3 labelled as low in the linguistic test. Classifier classifies 7 questions as high, 9 as average and 4 as low giving a score of 43 in the linguistic test. Total intelligence score for subject 3 is also 115 which maps to average intelligence.

Table 7 contains the detailed result of subject 3. In the memory test, subject 1 answered 8 questions perfectly, 10 questions with average label and remaining 2 were labelled as low. Whereas, the predictions made by the classifier gives 9 questions as high, 11 questions as average and not a single question as low. The total score for memory test is calculated to be 49. In arithmetic test, the subject's performed as follows. 6 questions were labelled as high, 11 were labelled as average and 3 were labelled as low. The classifier predicts 7 high, 8 average and 5

Table 5
Analysis of subject 1.

Question Number	Memory test			Arithmetic test			Linguistic test		
	Labelled output	Predicted output	Score	Labelled output	Predicted output	Score	Labelled output	Predicted output	Score
1	Low	Low	1	High	High	3	High	High	3
2	Average	Average	2	Average	Average	2	Average	Average	2
3	Low	Low	1	High	High	3	Average	Low	1
4	Low	Low	1	Low	High	3	Average	Average	2
5	Average	Average	2	Low	Low	1	High	High	3
6	Average	Average	2	Low	High	3	Average	Low	1
7	Average	Low	1	Average	Average	2	Low	Low	1
8	Average	Average	2	Low	Average	2	High	High	3
9	Average	Average	2	Average	Average	2	High	High	3
10	High	High	3	Average	High	3	Average	Low	1
11	Low	Low	1	Average	Average	2	Low	Low	1
12	Average	Low	1	Average	High	3	Average	Average	2
13	Low	Low	1	Low	Low	1	Low	Low	1
14	High	Low	1	Average	Average	2	Low	Average	2
15	High	High	3	Average	Low	1	Low	Low	1
16	High	High	3	Average	Average	2	Average	Average	2
17	Low	Low	1	High	High	3	Low	Low	1
18	Average	Low	1	Low	Low	1	High	High	3
19	Average	Average	2	Average	Average	2	Average	High	3
20	Low	High	3	Low	Low	1	High	High	3

low outputs. Subject's arithmetic score is found to be 42. In linguistic test, 6 questions were labelled as high, 5 were labelled as average and 9 were labelled as low. Classification's output has 7 questions predicted as high, 5 as average and 8 as low. The linguistic score is calculated to be 39, giving a total of 131 intelligence score which falls under average intelligence according to Fig. 9.

Subject 4′s results are given in Table 8. Subject's memory test performance was as follows. 6 questions were labelled high, 7 questions were labelled average and 7 were labelled low. The classifier however, predicts 7 questions as high, 6 as average and 7 as low, giving a total of 39 for memory test. In arithmetic test, student 4 had 4 questions labelled as high, 7 as average and 9 as low. The number of classifier's predictions are same as the number of class labels in the arithmetic test. Although few questions are predicted incorrectly. Subject's score for arithmetic test is 35. In the linguistic test, subject 4 gave 6 answers perfectly, 11 answers with one or two positions

correct and 3 answers absolutely wrong. The classifier predicts high for 10 questions, average for 8 questions and low for 2 questions. The linguistic score for subject 4 is 48. The final intelligence score for subject 4 is 112 which falls under average intelligence (Table 9).

4.3. Comparison with Wechsler Test

The subjects perform the Wechsler IQ test (Wechsler Test, 2019) i.e. the long IQ test which contains 50 questions with a time limit of 12 min. Every subject's first attempt is considered. Table 10 contains the results of the long IQ test for each subject. Subject 1 answered 25 out of 50 questions correctly which corresponds to IQ in range [107, 114] which falls under Middle High Average category. Subject 2 and subject 4 both answered 30 out of 50 questions correctly, scoring IQ in range [116, 125] which comes in High Average category. Subject 3 answered 38 out of 50 questions correctly which corresponds to IQ in [127, 134] and is categorized as Superior.

Table 6
Analysis of subject 2.

Question Number	Memory test			Arithmetic test			Linguistic test		
	Labelled output	Predicted output	Score	Labelled output	Predicted output	Score	Labelled output	Predicted output	Score
1	Average	Average	2	Average	Average	2	High	High	3
2	Average	Average	2	Low	Low	1	Average	High	3
3	Average	Average	2	Average	Average	2	Low	Low	1
4	Average	High	3	Average	Average	2	Average	Average	2
5	High	High	3	Low	Low	1	High	Average	2
6	Average	Low	1	Low	Low	1	Average	Average	2
7	Average	Average	2	Average	Average	2	High	High	3
8	Low	High	3	Average	Average	2	Average	Average	2
9	Low	Low	1	Average	High	3	Average	Average	2
10	High	High	3	Low	Low	1	Average	High	3
11	Average	Average	2	High	High	3	High	High	3
12	Average	Low	1	Average	Low	1	Average	Low	1
13	High	High	3	High	High	3	Low	Low	1
14	Low	Low	1	Average	Average	2	High	High	3
15	Low	Low	1	Average	Average	2	Low	Low	1
16	High	Low	1	Low	Low	1	High	High	3
17	Low	Low	1	Average	Low	1	Average	Average	2
18	Low	Low	1	High	High	3	Average	Average	2
19	High	High	3	Low	Low	1	Average	Average	2
20	Average	Low	1	High	Low	1	Average	Average	2

Table 7
Analysis of subject 3.

Question Number	Memory test			Arithmetic test			Linguistic test		
	Labelled output	Predicted output	Score	Labelled output	Predicted output	Score	Labelled output	Predicted output	Score
1	High	High	3	Average	Average	2	Low	Low	1
2	Average	Average	2	Low	Low	1	Low	Average	2
3	Average	Average	2	Average	Average	2	Average	Average	2
4	High	High	3	High	High	3	Low	Low	1
5	Average	High	3	High	High	3	High	High	3
6	High	High	3	Average	Average	2	High	High	3
7	High	High	3	High	High	3	Average	High	3
8	High	High	3	Average	Average	2	Low	Low	1
9	Average	Average	2	High	High	3	Low	Low	1
10	Low	Average	2	Average	Average	2	High	High	3
11	Average	Average	2	Average	Average	2	High	High	3
12	High	High	3	Low	Low	1	Average	Average	2
13	Average	Average	2	Average	Average	2	High	Low	1
14	Average	Average	2	Average	Average	2	Low	High	3
15	Average	Average	2	High	High	3	High	High	3
16	High	High	3	Average	Average	2	Low	Low	1
17	Average	Average	2	Average	Average	2	Average	Average	2
18	Low	Average	2	Average	Average	2	Average	Average	2
19	Average	Average	2	High	High	3	Low	Low	1
20	High	High	3	Low	Low	1	Low	Low	1

4.4. Statistical analysis

The SPSS software is used to carry out various statistical measurements. Mean and standard deviation is measured for the results obtained from EEG approach and from standard Wechsler test. Paired *t*-test (Tae Kyun Kim, 2015) between the results of EEG approach and the results of Wechsler IQ test is conducted.

- Mean intelligence score:
- EEG approach:

$$Mean = \frac{\sum Score\ of\ each\ subject}{Number\ of\ subjects} = \frac{115 + 115 + 131 + 112}{4} = 118.25$$

- Wechsler test:

$$Mean = \frac{\sum Score\ of\ each\ subject}{Number\ of\ subjects} = \frac{110.5 + 120.5 + 130.5 + 120.5}{4}$$
$$= 120.5$$

- Standard Deviation:
- EEG approach:

Table 9
Calculation of Intelligence score and the category of each Subject.

Subject	Memory score	Arithmetic score	Linguistic score	Intelligence score	Category
1	34	42	39	115	Average
2	37	35	43	115	Average
3	49	43	39	131	Average
4	39	35	48	112	Average

Table 8
Analysis of subject 4.

Question Number	Memory test			Arithmetic test			Linguistic test		
	Labelled output	Predicted output	Score	Labelled output	Predicted output	Score	Labelled output	Predicted output	Score
1	High	High	3	Average	Average	2	Average	Average	2
2	Low	Average	2	High	Low	1	High	High	3
3	Average	Average	2	Average	Average	2	Average	Average	2
4	Average	Average	2	Average	Average	2	High	High	3
5	High	High	3	Low	Low	1	Low	Low	1
6	Average	High	3	Low	Low	1	Average	High	3
7	Low	Low	1	Average	Average	2	High	High	3
8	High	High	3	Average	Average	2	Average	Average	2
9	Low	Low	1	Low	High	3	High	Average	2
10	Low	Low	1	Low	Low	1	Average	Average	2
11	Low	Low	1	High	High	3	High	High	3
12	Average	Low	1	Low	Low	1	Low	High	3
13	High	High	3	High	High	3	Average	Average	2
14	High	High	3	Average	Average	2	Average	Average	2
15	Low	Low	1	Average	Average	2	Average	High	3
16	Low	Low	1	Low	Low	1	Average	Average	2
17	Average	Average	2	Low	Low	1	High	High	3
18	High	High	3	High	High	3	Average	High	3
19	Average	Average	2	Low	Low	1	Low	Low	1
20	Average	Average	2	Low	Low	1	Average	High	3

Table 10
Welsher IQ test results.

Subject	Correct Answers	Minimum IQ	Maximum IQ	Category
1	25	107	114	Middle High Average
2	30	116	125	High Average
3	38	127	134	Superior
4	34	116	125	High Average

$$\sigma = \sqrt{\frac{\sum (Score\ of\ each\ subject\ -\ Mean)^2}{Number\ of\ subjects}} = 7.4624$$

- Wechsler test:

$$\sigma = \sqrt{\frac{\sum (Score\ of\ each\ subject\ -\ Mean)^2}{Number\ of\ subjects}} = 7.0710$$

- Paired t-test:

Paired *t*-test is performed between the results of Wechsler test and EEG approach to determine whether the mean difference between those two results is zero.

- Parameters:

$Significance\ level = \alpha = 0.05$

$Expected\ difference = \mu_0 = 0$

$Null\ hypothesis\ (H_0){:}\ \mu = \mu_0$

$Alternative\ hypothesis\ (H_1){:}\ \mu \neq \mu_0$

- Calculations:

$Mean\ difference = \bar{x}_d = -2.25$

$Standard\ deviation\ of\ difference = S_d = 5.85235$

$Skewness = 0.165883\ (potentially\ symmetrical)$

$$t = \frac{\bar{x}_d}{S_d \div \sqrt{n}} = \frac{-2.25}{5.85235 \div \sqrt{4}} = -0.768922$$

$$p - value = 2 \times P(x \leq t) = 0.497929$$

- Validations:

Since $p - value > \alpha$, H_0 is accepted.

The average of the EIQ scores minus Wechsler test IQ scores is considered to be equal to the μ_0. The test statistic $t = -0.768922$, is in the 95% critical value accepted range: $[-3.1824 : 3.1824]$

$\bar{x}_d = -2.25$, is in the 95% accepted range: $[-9.31 : 9.31]$

4.5. Spectral analysis

Fig. 14 shows the spectral analysis (Dressler et al., 2004) of EEG data obtained during the three different tests using spectopo.m (Spectopo.m, 2019) in EEGLAB (Delorme and Makeig, 2004) in MATLAB (MATLAB, 2017). The red coloured area corresponds to higher brain activity in that region of the brain, whereas blue coloured area corresponds to lower brain activity in that region of the brain. During memory test, the frontal lobe and temporal lobes show higher brain activity. During arithmetic test, the frontal lobe shows higher brain activity, whereas during linguistic test, the temporal lobe shows higher brain activity. This observation supports the facts mentioned in the setup section 3.2 about the involvement of various brain lobes with respect to various brain activities.

5. Conclusion

This paper focuses on developing a novel score scheme for measuring human intelligence. To achieve this goal BCI approach using EEG signal has been utilized. An experiment is performed using an EEG device, consisting of three tests: memory, arithmetic and linguistic. The EEG signals are converted into numerical attributes. For feature extraction, Wavelet Packet Transform was used and Daubechies wavelets have been applied for the same. For classification of the performance of each task Hierarchical Extreme Learning machine is implemented as classifier. The scoring for each task is determined by the predictions made using H-ELM classifier. The classifier achieved 80.00% of training accuracy and 73.33% testing accuracy. The aggregate score of each subject has been calculated and the subjects have been assigned to different categories according to the to the range as specified in Fig. 9.

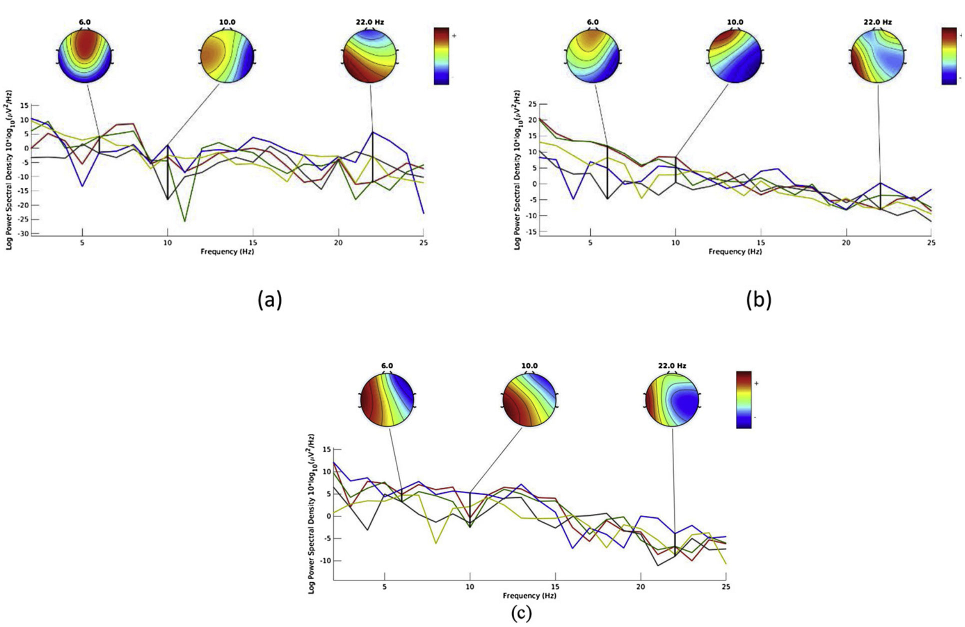

Fig. 14. Spectral analysis of EEG data during (a) Memory test, (b) Arithmetic test, (c) Linguistic test.

The Wechsler test has been answered by each subject and the corresponding scores have been displayed for the subjects. Comparison between Wechsler test scores and proposed EIQ scores show that the scores are in agreement with each other. The statistical analysis has been carried out on the results of both, Wechsler test and EIQ test, using SPSS software and the results of the analysis are displayed. The power spectral analysis of the EEG signal was performed and the results were displayed in Fig. 14. The advantages of the proposed EIQ score are that it covers intelligence over all three major domains - Memory, Arithmetic and Linguistics. Also, it measures the brain activity instead of final answers. This ensures that there is no luck factor involved as guessing the correct answer does not guarantee that EIQ score will increase.

References

Abdulkader, S.N., Atia, A., Mostafa, M.S.M., 2015. Brain computer interfacing: applications and challenges. Egypt. Inform. J. 16, 213–230.

Ahmed, Shah Aqueel, Rani, Dr.D.Elizabath, Sattar, Dr.Syed Abdul, 2012. Alpha activity in EEG and intelligence. Int. J. Adv. Inform. Technol. (IJAIT) 2 (February (1)).

American Electro-Encephalographic Society, 1994. Guideline thirteen: guidelines for standard electrode position nomenclature. J. Clin. Neurophysiol. 11, 111–113.

Beck, Amir, 2009. Marc Teboulle a fast iterative shrinkage-thresholding algorithm for linear inverse problems, society for industrial and applied mathematics. SIAM J. Imaging Sci. 2 (1), 183–202.

Binet, Alfred, 1916. New methods for the diagnosis of the intellectual level of subnormals. The development of intelligence in children: the Binet-Simon Scale. E.S. Kite (Trans.). Baltimore: williams & Wilkins. Pp. 37–90. Retrieved 10 July 2010. originally published as méthodes nouvelles pour le diagnostic du niveau intellectuel des anormaux. 1905. L'Année Psychologique 11, 191–244.

Bochner, S., Chandrasekharan, K., 1949. Fourier Transforms. Princeton University Press.

Brain products, 2019. http://www.brainproducts.com/.

Cao, Le-le, Huang, Wen-bing, Fu-chun, Sun., 2016. Building feature space of extreme learning machine with sparse denoising stacked-autoencoder. Neurocomput 174 (January), 60–71. https://doi.org/10.1016/j.neucom.2015.02.096. PA2016.

Daubechies, I., 1992. Ten Lectures on Wavelets. SIAM p. 194.

Delorme, A., Makeig, S., 2004. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics. J. Neurosci. Methods 134, 9–21. https://sccn.ucsd.edu/eeglab/index.php.

Ding, Jian-Jiun, 2008. Time-Frequency Analysis and Wavelet Transform.

Dressler, O., Schneider, G., Stockmanns, G., Kochs, E.F., 2004. Awareness and the EEG power spectrum: analysis of frequencies. Br. J. Anaesth. 93 (6), 806–809.

EEG Head picture, 2019. https://www.123rf.com/profile_hakaba/new?mediapopup=36374069.

Goodfellow, Ian, Bengio, Yoshua, Courville, Aaaron, 2016. Deep Learning. MIT Press p. 196. ISBN 9780262035613.

Graps, A., 1995. Introduction to wavelets. Ieee Comput. Sci. Eng. 2, 50–61.

Grossmann, A., Morlet, J., 1984. Decomposition of Hardy functions into square integrable wavelets of constant shape. Soc. Int. Am. Math. (SIAM), J. Math. Analys. 15, 723–736.

Guang-Bin, Huang, Qin-Yu, Zhu, Chee-Kheong, Siew, 2006. Extreme learning machine: theory and applications. Neurocomputing. 70 (1), 489–501.

Homan, R.W., Herman, J., Purdy, P., 1987. Cerebral location of international 10–20 system electrode placement. Electroencephalogr. Clin. Neurophysiol. 66, 376–382.

Huang, Guang-Bin, Zhu, Qin-Yu, Siew, Chee-Kheong, 2004. Extreme learning machine: a new learning scheme of feedforward neural networks. 2004 IEEE Int. Joint Confer. Neural Net.

IBM Corp. Released 2017. IBM SPSS Statistics for Windows, Version 25.0. Armonk, NY: IBM Corp. https://www.ibm.com/in-en/analytics/spss-statistics-software.

Image of 10-20 International System, 2019. http://eegatlas-online.com/index.php/en/montages.

Image of brain, 2019. http://animalia-life.club/other/human-brain-clipart-black-and-white.html.

Imotions, 2016. EEG (Electroencephalography): The Complete Pocket Guide.

Shen, Jianhong (Jackie), Strang, Gilbert, 1998. Asymptotics of daubechies filters, scaling functions, and wavelets. Applied and Computational Harmonic Analysis 5 (3).

Leinweber, D.J., 2007. Stupid data miner tricks. J. Invest. 16, 15–22. https://doi.org/10.3905/joi.2007.681820.

Lodish, H., Berk, A., Zipursky, S.L., et al., 2000. In: Freeman, W.H. (Ed.), Molecular Cell Biology, 4th ed. W. H. Freeman and Company, New York (2017).

MATLAB, 2017. Version R2017a. The MathWorks Inc., Natick, Massachusetts. https://www.mathworks.com/products/matlab.html.

Moore, E.H., 1920. On the reciprocal of the general algebraic matrix. Bull. Am. Math. Soc. 26 (9), 394–395. https://doi.org/10.1090/S0002-9904-1920-03322-7.

Mustafa, Mahfuzah, Taib, Mohd Nasir, Lias, Sahrim, Murat, Zunairah Hj, 2013. Norizam sulaiman EEG spectrogram classification employing ANN for IQ application. The International Conference on Technological Advances in Electrical, Electronics and Computer Engineering (TAEECE).

Niedermeyer, E., da Silva, F.L., 2004. Electroencephalography: Basic Principles, Clinical Applications, and Related Fields. Lippincott Williams & Wilkins ISBN 978-0-7817-5126-1.

Powers, David M.W., 2011. Evaluation: from precision, recall and F-Measure to ROC, informedness, markedness & correlation. J. Mach. Learn. Technol. 2 (1), 37–63.

Reader, Simon M., Hager, Y., Laland, K.N., 2011. The evolution of primate general and cultural intelligence. Philos. Trans. Biol. Sci. 366 (1567), 1017–1027.

Shenoi, Belle A., 2006. Introduction to Digital Signal Processing and Filter Design. John Wiley and Sons p. 120. ISBN 978-0-471-46482-2.

Singh, Nivedita, Ansari, M.A., 2016. Extraction of feature based of wavelet packet transform. 2016. 1st India International Conference on Information Processing.

Spectopo.m, 2019. Scott Makeig & Marissa Westerfield, SCCN/INC/UCSD, La Jolla, 3/01. https://sccn.ucsd.edu/~arno/eeglab/auto/spectopo.html.

Robert J. Sternberg, 2019. Human Intelligence, Britannica Encyclopedia. https:/www.britannica.com/science/human-intelligence-psychology.

Kim, Tae Kyun, 2015. T test as a parametric statistic. Korean J. Anesthesiol. 68 (December (6)), 540–546.

Tang, Jiexiong, Deng, Chenwei, Huang, Guang-Bin, 2015. Extreme learning machine for multilayer perceptron. IEEE Trans. Neural Netw. Learn. Syst. 4, 809–821.

Thatcher, R.W., Palmero-Soler, E., North, D.M., Biver, C.J., 2016. Intelligence and Eeg Measures of Information Flow: Efficiency and Homeostatic Neuroplasticity.

Wavedec (MATLAB), 2019. http://in.mathworks.com/help/wavelet/ref/wavedec.html.

Wechsler, D., 1944. The Measurement of Adult Intelligence. Williams & Wilkins, Baltimore.

Wechsler Test, 2019. https://wechslertest.com/.